The atomic orbitals of the topological atom

Eloy Ramos-Cordoba, Pedro Salvador, and István Mayer

Additional information on J. Chem. Phys.

The atomic orbitals of the topological atom

Eloy Ramos-Cordoba,[1] Pedro Salvador,[1,a)] and István Mayer[2]

[1]*Institute of Computational Chemistry and Catalysis and Department of Chemistry, University of Girona, 17071 Girona, Spain*

[2]*Institute of Organic Chemistry, Research Centre for Natural Sciences, Hungarian Academy of Sciences, P.O. Box 17, H-1525 Budapest, Hungary*

The effective atomic orbitals have been realized in the framework of Bader's atoms in molecules theory for a general wavefunction. This formalism can be used to retrieve from any type of calculation a proper set of orthonormalized numerical atomic orbitals, with occupation numbers that sum up to the respective Quantum Theory of Atoms in Molecules (QTAIM) atomic populations. Experience shows that only a limited number of effective atomic orbitals exhibit significant occupation numbers. These correspond to atomic hybrids that closely resemble the core and valence shells of the atom. The occupation numbers of the remaining effective orbitals are almost negligible, except for atoms with hypervalent character. In addition, the molecular orbitals of a calculation can be exactly expressed as a linear combination of this orthonormalized set of numerical atomic orbitals, and the Mulliken population analysis carried out on this basis set exactly reproduces the original QTAIM atomic populations of the atoms. Approximate expansion of the molecular orbitals over a much reduced set of orthogonal atomic basis functions can also be accomplished to a very good accuracy with a singular value decomposition procedure. © *2013 AIP Publishing LLC*. [http://dx.doi.org/10.1063/1.4807775]

INTRODUCTION

The concept of *atom in a molecule* has always craved for a proper definition. However, we are lacking a single, unambiguous one. Instead, over the last decades a number of schemes or formalisms have been devised to identify the atom *within* a molecule.[1–10] Probably, any of such schemes considers the nucleus as part of the atom, so the differences always arise in how the electron population distributed in the physical space (or in the Hilbert-space) is subdivided into atomic shares.

Within the LCAO approach, the atom may be identified with the subspace of the basis functions attached to it. Such approach leads to the so-called Hilbert-space analyses,[11] such as the classical Mulliken[1] or Löwdin[2] population analysis of the density. Despite their simplicity, Hilbert-space analyses have been criticized by their restricted applicability (the use of atom-centered basis functions is necessary) and their notable basis set dependence. The latter represents a true flaw when using extended basis sets including diffuse functions.[12] Alternative population analyses based upon occupation numbers[13,14] carried out onto an AO basis set different from the extended one minimize the basis set dependence.

A different strategy is to subdivide the physical three-dimensional (3D) space into atomic regions or domains, which represent (together with the nucleus) the atom. These domains may be defined disjoint, like in Bader's atoms in molecule theory[4] (often referred to as QTAIM—"quantum theory of atoms in molecules"), or may be allowed to overlap, like in the different flavors of "fuzzy" atoms.[3,5–10]

The 3D space formalism represents a perfect counterpart of the Hilbert-space analysis. Indeed, by introducing a proper mapping,[15,16] one can find a one-to-one correspondence between the expressions of quantities such as atomic populations and bond orders,[5] energy components,[17–19] or local spins,[20] obtained in these two frameworks. Of course, the actual values of these quantities derived from one or another formalism differ. Moreover, the QTAIM analysis may yield domains with so-called non-nuclear attractors, which correspond to regions of the space with no nucleus associated. Often, the appearance of a non-nuclear attractor is an artifact of the basis set applied, such as in the case of acetylene.[21,22] In such cases, it may be worth to use another partitioning of the space.

When looking for an *atom in a molecule*, obviously we are not merely interested in a subdivision of the 3D space into atomic volumes, but rather in assigning different physical quantities to the individual atoms (or their groups). Any physical quantity is expressed as the expectation value of one-electron (two-electron) operators, and it can also be written in terms of the integral over the space of the appropriate one-electron (two-electron) density functions. Hence, the subdivision of the space into atomic regions naturally leads to the decomposition of different physical quantities into atomic (diatomic) terms that can be considered the *effective* atomic (or diatomic) values of that physical quantity within the molecule. This has been extensively accomplished within the QTAIM framework, partly due to the special properties provided by the zero-flux condition, such as the local fulfillment of the virial theorem.[4] However, it has been shown that one can obtain quantities such as atomic populations and valences,[5] energy components,[17–19] or local spins[20] for essentially any atom in molecule definition. Similarly, the atom in

a)E-mail: pedro.salvador@udg.edu

the molecule can also be characterized from the analysis of the density matrix.[23–25]

Probably, the most appropriate entities that serve to characterize the state of the atom within the molecule are the so-called effective atomic orbitals ("effective AOs").[26–31] In this approach, one obtains for each atom a set of orthogonal atomic hybrids and their respective occupation numbers usually adding up to the *net* population of the atom. A remarkable feature of this scheme is that, for a given molecular system, irrespective of the basis set size applied in the calculation, one in practice obtains the same number of significantly populated atomic hybrids. This number has always been found equal to the number of the orbitals in a classical minimal basis, which means that in this manner one obtains "effective minimal basis sets" by performing *a posteriori* analyses of the calculations (special provisions should be made in the case of hypervalent systems). This happens to be the case even if no atom-centered basis functions are used at all.[31] These atomic hybrids closely mimic the core and valence shells of the atom, as anticipated on the basis of classical notions of electron configuration of the atom/fragment within the molecule.

This formalism was first introduced nearly two decades ago in the framework of Hilbert-space analysis,[27] and then it was generalized for an arbitrary Hermitian bilinear "localization functional."[28] It was applied to the case of "fuzzy" atoms a few years ago.[30] It had also been formulated in the context of Bader's QTAIM,[29] but until now that version was never actually realized.

When the atoms are associated with non-overlapping domains, as is the case of Bader's QTAIM, the "effective AOs" have special properties that make them very appealing from both conceptual and practical points of view, as will be shown later on. The first feature is that the "effective AOs" associated to different atoms are also orthogonal, as a direct consequence of not sharing at all the physical space. Also, because the atoms are not allowed to overlap, the atom's net and gross populations are equal. In other words, the sum of the occupation numbers of the "effective AOs" of a given atom is equal to its atomic population.

It is fair to note that the formalism of the domain-averaged Fermi hole analysis[32,33] (DAFH) also produces orbital functions in the framework of QTAIM (or other AIM schemes). With the DAFH analysis, one first obtains a set of domain orbitals and orbitals occupancies, that have their origin in the average of the exchange-correlation density over a space domain, typically the union of several atomic domains. These orbitals are then localized via a non-unitary isopycnic transformation,[34] leading to a new set of objects that are interpreted in terms of bonding orbitals, lone pairs, and dangling valences. Even though their origin is different, the DAFH and "effective AO" analyses share some similarities, particularly when the DAFH analysis is carried out over a single atomic domain. In fact, in the restricted single-determinant case the orbitals before isopycnic localization, sometimes denoted as domain natural orbitals,[33] are exactly the same as the original "effective AOs."[29]

The special features of the "effective AOs" in the framework of QTAIM (or any disjoint partition of the space in general) suggests that they could be used as (numerical) atomic basin-centered orthogonal basis set, in which the actual molecular orbitals of the molecule can be expanded. Furthermore, since the "effective AOs" can be obtained even if no atomic basis functions are used at all[31] (e.g., plane wave calculations), this formalism can be used to actually retrieve from such type of calculations a proper set of orthogonal atomic basis functions.

Of course, the fact that the molecular orbitals can be expressed as a linear combination of an alternative set of basis functions, even if it is orthonormal and has a reduced number of functions, adds no special chemical relevance *per se*. But, as we show in the "Theoretical Methods" section, in this particular basis, the Hilbert-space and 3D space analyses turn out to be fully equivalent numerically. Thus, the Mulliken population analysis of the density carried out on the basis of "effective AOs" exactly reproduces the original QTAIM atomic populations of the atoms.

In other words, these "effective AOs" appear to be the genuine atomic orbitals of Bader's theory, perhaps one of the few ingredients missing in QTAIM's toolbox.

THEORETICAL METHODS

Let us consider a system with n orthonormalized doubly occupied orbitals $\varphi_i(\vec{r})$, $i = 1, 2\ldots, n$, and the division of the 3D space into N_{at} disjunct atomic domains Ω_A defined, e.g., by Bader's "topological" QTAIM method. Let us for each atom A ($A = 1, 2, \ldots, N_{at}$) form the $n \times n$ Hermitian matrix $\mathbf{Q}^A$ with the elements

$$Q_{ij}^A = \int_{\Omega_A} \varphi_i^*(\vec{r})\varphi_j(\vec{r})dv. \tag{1}$$

Matrix $\mathbf{Q}^A$ is essentially the "atomic overlap matrix" in the basis of the MOs φ_i.

Furthermore, for each atom A we define the "intraatomic" part φ_i^A of every MO φ_i as

$$\varphi_i^A(\vec{r}) = \begin{cases} \varphi_i(\vec{r}) \text{ if } \vec{r} \in \Omega_A; \\ 0 \quad \text{ if } \vec{r} \notin \Omega_A. \end{cases} \tag{2}$$

Thus,

$$Q_{ij}^A = \langle \varphi_i^A | \varphi_j^A \rangle, \tag{3}$$

i.e., $\mathbf{Q}^A$ is the overlap matrix of the orbitals φ_i^A.

Owing to the disjunct character of the atomic domains, one obviously has

$$\varphi_i(\vec{r}) = \sum_A^{N_{at}} \varphi_i^A(\vec{r}). \tag{4}$$

We diagonalize the Hermitian matrix $\mathbf{Q}^A$ by the unitary matrix $\mathbf{U}^A$

$$\mathbf{U}^{A\dagger}\mathbf{Q}^A\mathbf{U}^A = \Lambda^A = diag\{\lambda_i^A\}. \tag{5}$$

It can be shown that every $\lambda_i^A \geq 0$, as is the case for every overlap matrix.

For each atom A, we can define n_A ($n_A \leq n$) "effective atomic orbitals" $\chi_\mu^A(\vec{r})$ as linear combinations of the

"intraatomic" parts $\varphi_i^A(\vec{r})$ of the MOs as

$$\chi_\mu^A(\vec{r}) = \frac{1}{\sqrt{\lambda_\mu^A}} \sum_{i=1}^n U_{i\mu}^A \varphi_i^A(\vec{r}) ; \quad \mu = 1, 2, \ldots, n_A, \quad (6)$$

where n_A is the number of non-zero eigenvalues λ_i^A.

One may consider rewriting the original MOs as linear combinations of the set of "effective AOs" with non-zero eigenvalue of different atoms, with the appealing result that the atomic populations calculated by the Hilbert-space analysis in the basis of the "effective AOs" are **equal** to those obtained by the 3D AIM analysis. The proof is given in Appendix A. Similarly, it can be seen that the classical Wiberg bond orders index calculated on the orthogonal basis of "effective AOs" exactly coincides with the bond orders[15] or (delocalization indices[35]).

In practice, the dimension of the "effective AO" basis has nothing to do with the dimension of the LCAO basis (if any) used in the original calculation. Equation (6) indicates that the maximum number of "effective AOs" that can be obtained is $N_{at} \times n$, which can be both less or more than the total number of the LCAO basis functions. Indeed, the proof provided in Appendix A considers up to n "effective AOs" per atom.

Moreover, experience shows that the number of "effective AOs" with significant occupation numbers on each atom is limited, and typically much smaller than the number n of the doubly occupied orbitals in the whole molecule. The remaining "effective AOs" with very small occupation numbers should have a marginal significance. Thus, one may think over expressing the molecular orbitals, to a good approximation as a linear combination of a (numerical) atomic basis set build up from a subset of the "effective AOs" (selected by an occupation number criterion), in an expansion similar to that of Eq. (A3) but with $n_A < n$. Since only a limited number (compared to the number of basis functions) of "effective AOs" typically exhibit significant occupation numbers, each molecular orbital could be expressed as a linear combination of a much reduced set of orthogonal basis functions with distinct atomic character, which may permit a much simpler analysis in chemical terms, specially when the MOs are expanded over an extended AO basis.

In order to obtain the new LCAO coefficients, one can make use of the singular value decomposition (SVD) technique to perform a pairing between the set of "effective AOs" from one side, and the set of doubly occupied MOs from the other. The procedure is described in Appendix B. In the Results and Discussion section, we will illustrate numerically how by this SVD process one can produce highly accurate atomic populations using a very much reduced set of "effective AOs." That is, the MOs can be expanded to a good approximation in terms of a minimal basis of "effective AOs."

Finally, it is worth to note that the scheme described above is not only applicable in the single-determinant closed-shell case. As noted by one of us,[28] the "effective AOs" of atom A can also be obtained from the diagonalization of the matrix $\mathbf{PS}^A$, where $\mathbf{P}$ is the LCAO density matrix and $\mathbf{S}^A$ is the atomic overlap matrix in the actual (AO or MO) basis. (In the single determinant case, if the dimension of the atomic basis applied is greater than the number of the occupied or-

bitals, there will be an appropriate number of strictly zero eigenvalues.)

This alternative also permits the straightforward generalization of the method both to the unrestricted single-determinant case and the correlated level, from which the $\mathbf{P}$ matrix is readily available. In that case, the SVD method can also be applied with some necessary adjustments, as indicated in Appendix B.

COMPUTATIONAL DETAILS

We have obtained the "effective AOs" in the framework of QTAIM for a series of molecules for illustrative purposes. The analysis has been performed by our program apost-3D.[36] The program includes a version of the grid-based scheme proposed by Rodríguez *et al.*[37] to integrate over atomic domains without the explicit calculation of the zero-flux surface. For the orbital plots, we have generated relatively large cubic atom-centered grids of $60 \times 60 \times 60$ points. The truncated nature of these orbitals makes rather difficult to obtain high-quality plots. An isosurface value of 0.15 has been used for all plots.

The *ab initio* calculations have been carried out with GAUSSIAN 03[38] program, employing the B3LYP density functional combined with cc-pVTZ basis set, unless otherwise indicated. For the triradical nitrene of Figure 6, we have also obtained the wavefunction at the complete active space self-consistent field (CASSCF) level of theory with an active space composed by 9 electrons and 9 orbitals. The active space includes the six orbitals of the π system of the ring, the σ orbital of the radical carbon, and the σ and π orbitals on the nitrogen atom. All calculations have been carried out at the geometrical structure of the molecules optimized at the current level of theory.

RESULTS AND DISCUSSION

Figure 1 depicts the "effective AOs" obtained for the C atom of methane. The corresponding occupation number is indicated below each orbital plot. As in methane there are only five doubly occupied MOs (in the single-determinant description), one can obtain up to five "effective AOs" for each atom. Thus, in this case there appear no weakly occupied ones on the carbon.

The first orbital can be clearly identified as a core $1s$-type orbital with an occupation number of 2.000. The next "effective AO" is a $2s$-type orbital, with an occupation number
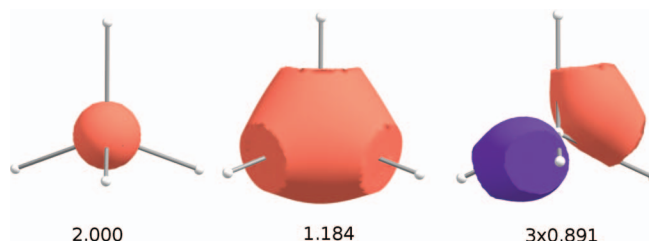

FIG. 1. "Effective AOs" for C atom in methane and their occupation numbers.
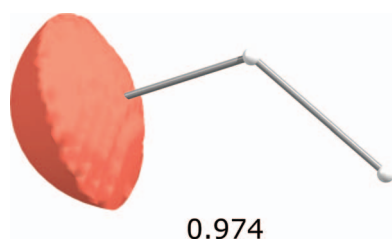
FIG. 2. "Effective AO" for H atom in methane and its occupation number.

of 1.189. The last plot of Figure 1 corresponds to one of the three degenerated (in the present context we use this term to denote orbitals with the same occupation numbers) $2p$-type atomic orbitals.

The main feature of the "effective AOs" is that they are completely confined within the corresponding atomic domain, i.e., in this case they are strictly cut at the boundary limit of the Bader atom. We have chosen a relatively small isocontour value on the plots to highlight this feature. The occupation numbers are typically close to 2 for core orbitals or lone pairs, while hybrids that are involved in bonding (valence "effective AOs") present occupations that oscillate around 1, depending upon how the atom is polarized by the presence of the neighboring atoms. The most electronegative atoms exhibit higher occupation numbers than the less electronegative ones. For instance, for the CF_4 molecule the "effective AOs" of the C atom are strikingly similar to those of Figure 1 (only smaller because the volume of the C atom is reduced), and only the occupation numbers differ. In this case, the occupations of the $2s$- and $2p$-type orbitals decrease to 0.444 and 0.296, respectively.

In the case of the H atoms of methane, only the truncated s-type orbital shown in Figure 2 has a significant occupation. There are two additional "effective AOs" of p-type with much lower occupations (less than 0.03, not shown).

The polarization of the H-atoms is already included in the most populated effective AOs which (within the atomic domain) need not to be pure s-functions. Similarly for carbon atoms, although the number of effective AOs is the same as in a minimal sp basis, the form of the orbitals may slightly deviate from the ideal s or p, and reflect, therefore, the polarization effects. The truncation of the orbitals, of course, introduces another, much bigger, "polarization."

The "effective AOs" with significant occupation numbers of the C, N, and O atoms of fulminic acid (HCNO) are shown in Figure 3. Again, five hybrids with non-negligible occupation numbers are obtained for each atom, associated to their core and valence orbitals. The orbitals of all three atoms exhibit similar shapes, and the occupation numbers of analogous orbitals increase from C to O, following the increase of electronegativity. One can identify a $1s$-type core orbital, with an occupation close to 2 in all cases. A somewhat distorted $2s$-type orbital is also present in all atoms. Remarkably, a similar orbital was also obtained for the C atom of the acetone from a wavefunction expressed in terms of plane waves and in the framework of the "fuzzy" atoms.[31] The occupation number in the case of the oxygen atom is almost 2, as the atomic domain has no boundary on the direction opposite to the nitrogen atom. Both carbon and nitrogen exhibit two boundaries, so the corresponding hybrids are truncated from both sides. A pair of degenerate $2p$-type orbitals on each atom are involved on the π system of this molecule. One can observe an apparent correspondence between the hybrids on carbon and nitrogen atoms. Their respective occupation numbers reflect that the atomic boundary between them is shifted towards the carbon atom. The two pairs of hybrids seem like "halves" (with different shares) of a common C–N π bond. However, it is worth to recall that they originate from two independent calculations. Similar complementarity is also observed between the two hybridized σ-type $2p$-orbitals along the internuclear axis of nitrogen and oxygen atoms, with occupations of 0.944 and 1.091, respectively.

The occupation number of the sixth "effective AO" of the carbon atom is essentially zero (smaller than 0.01). In the oxygen atom it is 0.03, but in the case of the nitrogen atom there
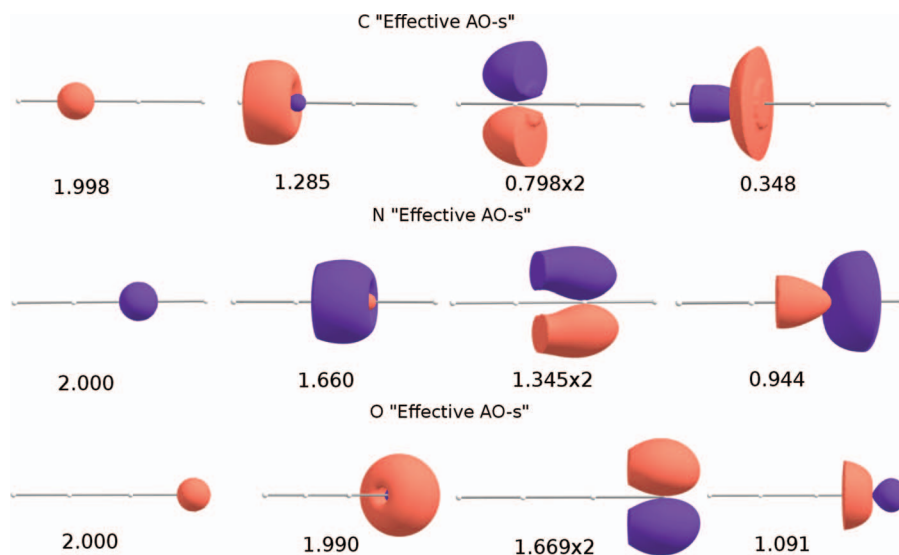

FIG. 3. "Effective AOs" and occupation numbers for the heavy atoms in fulminic acid.
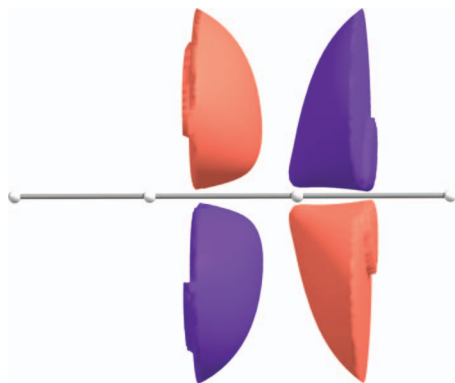
FIG. 4. Weakly occupied "effective AO" of the nitrogen atom of fulminic acid.

is a pair of degenerate "effective AOs" with a more significant occupation number of 0.133. This is a distinct manifestation of the hypervalent character of nitrogen in this molecule. In fact, Karadakov *et al.*[39] already discussed this phenomenon in this molecule with a rather involved analysis combining valence-bond and CASSCF calculations.

The shape of this weakly occupied hybrid is shown in Figure 4. It corresponds to a strongly polarized (its centroid appears displaced from the nuclear position) *d*-type orbital in a plane containing the four atoms, and with the proper orientation in order to be involved in the pi-bonding system of the molecule. Even though the occupation number is small, its participation is necessary to explain how this nitrogen atom can be involved in (formally) more than four covalent bonds (three with the carbon atom and two with the oxygen atom).

Another example of hypervalent behaviour is given by the sulfur atoms in the series $CH_3SO_xCH_3$, x = 0, 1, 2. In this case, occupation numbers of the "effective AOs" of the sulfur atoms are displayed in Figure 5.

The first observation is that the sulfur atoms exhibit in all cases five "effective AOs" with occupation number equal to 2, that correspond to the doubly occupied 1*s*, 2*s*, 2*p* inner shell. Inspection of the curve of the sulfur of dimethyl sulfide

(squares) reveals the presence of two more almost doubly occupied hybrids, corresponding to lone pairs. The first has a strong 3*s* character while the second one is of 3*p*-type, perpendicular to the plane formed by the sulfur and the two carbon atoms. The next two hybrids have an occupation close to 1 and are oriented towards the carbon atoms, indicating that they are involved in the sulfur-carbon σ bonds. The occupations from the ninth hybrid drop to a very small value (∼0.05). Hence, the number of "effective AOs" with significant occupation number is again equal to the number of orbitals in the classical minimal basis (nine orbitals for 3rd row elements).

The curve of dimethyl sulfoxide (diamonds) indicates that there is only one lone pair. The 3*p*-type hybrid that in the case of dimethyl sulfide corresponded to a lone pair, now appears oriented towards the oxygen atom with occupation number of 0.472. Such decrease in the occupation is caused by the stronger electronegativity of the oxygen atom. The shape of the two remaining hybrids involved in the sulfur-carbon σ bonds is changed by the presence of the oxygen atom, but their occupation numbers are remarkably similar to those obtained in the case of dimethyl sulfide. In the inset of Figure 5, one can see that in this case there appear two more "effective AOs" with occupation numbers close to 0.10. These hybrids are of *d*-type and contribute to provide a slight π character to the sulfur-oxygen bonding. Similar to the case of the central N atom in fulminic acid, the participation of these *d*-orbitals is necessary to account for the formal double bond between the S and O atoms. Finally, for the sulfur atom of the dimethyl sulfone molecule there appear no doubly occupied valence hybrids. Instead, the two hybrids that are to be involved in the bonding with the oxygen atoms now exhibit occupations of 0.482 and 0.381. Moreover, there is a larger number of "effective AOs" with significant occupation numbers beyond the valence shell. The inset of the figure shows that there are essentially four hybrids of *d*-type that have a small but non-negligible contribution. The occupation number of the fifth *d*-type orbital (0.055) is very similar to that of the ninth hybrid in the case of dimethyl sulfide molecule. Such a behaviour is characteristic for the effective AOs of the hypervalent sulfur in the framework of the Hilbert space analysis, too; for an early example see Ref. 27.

For illustrative purposes, we also consider a singular nitrene triradical[40] molecule with a quartet ground state. We have obtained the "effective AOs" from the wavefunctions calculated for the quartet ground state at the CASSCF(9,9) and B3LYP levels of theory. The respective occupation numbers for selected atoms (for their numeration see Fig. 6) are gathered on Table I. The pictures of the orbitals do not differ significantly and are not shown.
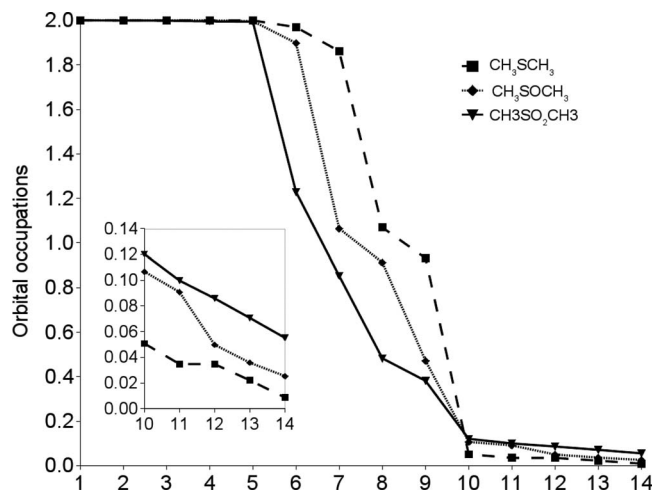

FIG. 5. Occupation numbers (in descending order) of the "effective AOs" of the sulfur atom of dimethyl sulfide, dimethyl sulfoxide, and dimethyl sulfone molecules.
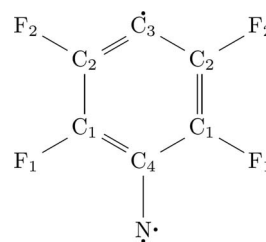

FIG. 6. Structure of the quartet fluorophenyl nitrene triradical.

TABLE I. Occupation numbers (larger than 0.01) of selected atoms of the nitrene triradical (see Fig. 6) obtained at the B3LYP and CASSCF(9,9) (in parentheses) levels of theory.

Orbital	C_1	C_3	C_4	F_1	N
$1s$	1.999 (1.999)	2.000 (2.000)	2.000 (2.000)	2.000 (2.000)	2.000 (2.000)
$2s$	1.131 (1.133)	1.295 (1.265)	1.074 (1.053)	1.998 (1.999)	1.991 (1.995)
$2p^a$	0.916 (0.899)	0.969 (0.960)	0.862 (0.819)	1.914 (1.951)	1.144 (1.121)
$2p^b$	0.845 (0.861)	0.865 (0.857)	0.793 (0.784)	1.942 (1.959)	0.983 (0.982)
$2p^c$	0.342 (0.283)	0.769 (0.779)	0.584 (0.486)	1.639 (1.702)	1.349 (1.461)
d-type	0.034 (0.036)	0.039 (0.037)	0.032 (0.030)	0.061 (0.069)	0.034 (0.048)
d-type	0.030 (0.026)	0.018 (0.016)	0.029 (0.028)	0.044 (0.050)	0.032 (0.044)
d-type	0.027 (0.023)	0.013 (0.015)	0.027 (0.020)	0.027 (0.026)	0.012 (0.014)
d-type	0.025 (0.021)	0.011 (0.010)	0.017 (0.013)		

[a]Out-of-plane.
[b]In-plane.
[c]In-plane towards the center of the ring.

The occupation numbers obtained with the two methods are strikingly similar, not only between the strongly occupied "effective AOs" but also between the weakly occupied ones. The largest difference in the occupations is only about 0.10, found for a $2p$ hybrids of the N and C_6 atoms. This indicates a slight change in the polarity of the C_4–N bond going from B3LYP to CASSCF. Each of the atoms contributes to the π system with roughly one electron except for F atoms, in which the $2p$ "effective AOs" is essentially doubly occupied (1.914). The $2p$ hybrids directed towards the center of the ring exhibit different occupation numbers from one atom to another. These "effective AOs" are the ones that are involved in forming the σ-type C–F and C–N bonds, and the radical center in C_3. Thus, the occupation numbers of the "effective AOs" for C_4 and N (0.584 and 1.349) and for C_1 and F (0.342 and 1.639) exhibit a good correspondence. That close parallel between the unrestricted B3LYP and CASSCF results may be understood if one takes into account that the CASSCF method accounts for valence correlation only. The use of a method which is able to account for the (basically intra-atomic) angular correlation would perhaps make the results more complicated – but hardly lead to a different overall picture.

The picture of Fig. 6 where one unpaired electron sits on the C_3 atom and two unpaired electrons sit on the N atom has no straight correspondence with the "effective AO" analysis given in Table I. In order to locate individual electrons (of a given spin), one can obtain the "effective AOs" from the alpha and beta densities separately.[41]

TABLE II. Occupation numbers of the valence "effective AOs" of the radical centers of the nitrene triradical (see Fig. 6) obtained at the B3LYP and CASSCF(9,9) (in parentheses) levels of theory.

Orbital	C_3 alpha	C_3 beta	N alpha	N beta
$2s$	0.528 (0.506)	0.577 (0.566)	0.996 (0.997)	0.996 (0.997)
$2p^a$	0.613 (0.651)	0.354 (0.309)	0.856 (0.854)	0.314 (0.287)
$2p^b$	0.431 (0.419)	0.415 (0.419)	0.935 (0.952)	0.000 (0.000)
$2p^c$	0.912 (0.936)	0.000 (0.000)	0.697 (0.723)	0.675 (0.723)

[a]Out-of-plane.
[b]In-plane.
[c]In-plane towards the center of the ring.

The respective occupation numbers for the two radical centers are collected on Table II. Note that the values of Table I are not exactly expressed as the sum of the occupation numbers of the alpha and beta contributions, as the "effective AOs" for each spin case are different. The differences in the occupation numbers between B3LYP and CASSCF are again minimal (less than 0.05). The radical center C_3 is characterized by an alpha "effective AO" with occupation close to 1, pointing outside the center of the ring. For the remaining "effective AOs" of this atom, the alpha and beta occupation numbers are similar, except for the out-of-plane hybrids forming the π system of the molecule. In that case, the occupation number of the alpha part almost doubles that of the beta. The picture of the other radical center (N atom) is very similar. The occupation numbers are in general larger than for the C_3 atom, due to the more electronegative character of nitrogen. There is a singly occupied "effective AO" of the alpha part, that lies in the molecular plane, perpendicular to the C_4–N bond. But there is no singly occupied $2p$ orbital perpendicular to the molecular plane. Instead, the occupation numbers of these alpha and beta "effective AOs" are not too different from those obtained for C_3, and as a matter of fact, for the rest of C atoms. Thus, the chemical picture that emerges from the analysis of the "effective AOs" is that there are two σ-type radical centers located on C_3 and N atoms, but there is a completely delocalized unpaired π-electron, as it is expected for such a benzonoid species. This is in agreement with the picture obtained from a local spin analysis.[42]

Finally, we illustrate how by the SVD procedure one can expand the MOs to a very good accuracy using a reduced subset of "effective AOs." We have applied the method for the alanine molecule computed with two different basis sets, namely, the 6-31++G** and the cc-pVTZ. The results are collected on Table III. We have used three different thresholds for the occupation numbers to select the subset of "effective AOs." For the cc-pVTZ basis set, up to 119 "effective AOs" have occupation number above a threshold of 0.001. This number decreases to just 71 for a threshold of 0.01, and to just 37 for a threshold of 0.1. Indeed, for this molecule the minimal basis set ($1s$ for H atoms and $1s2s2p$ for the rest) includes 37 atomic orbitals. Once the "effective AOs" are selected, the SVD procedure is used to obtain the LCAO

TABLE III. Accuracy of the SVD procedure for Ala molecule with two basis sets (number of basis functions in parentheses). N_e indicates the number of effective atomic orbitals with occupation number above the threshold. $\bar{\delta}$ and δ_{max} are the average and maximum error in the atomic population values after the SVD procedure.

	6-31++G(d,p) (156)	cc-pVTZ (278)
Occ. number >0.001		
N_e	114	119
$\bar{\delta}$	8.9×10^{-4}	8.8×10^{-4}
δ_{max}	1.9×10^{-3}	1.6×10^{-3}
Occ. number >0.01		
N_e	69	71
$\bar{\delta}$	7.6×10^{-3}	7.3×10^{-3}
δ_{max}	1.7×10^{-2}	1.6×10^{-2}
Occ. number >0.1		
N_e	37	37
$\bar{\delta}$	3.2×10^{-2}	3.1×10^{-2}
δ_{max}	9.8×10^{-2}	8.9×10^{-2}

coefficients of the MOs in the new numerical atomic basis. Then, the atomic populations are calculated with classical Mulliken population analysis on this basis set.

The $\bar{\delta}$ and δ_{max} on Table III account for the average and maximum error in the recomputed atomic populations, with respect to those originally obtained by the 3D-space QTAIM method. It is worth to recall that the total number of electrons is conserved after this transformation. Thus, any lack of flexibility of the numerical basis set to expand the MOs will be translated in fluctuations of the atomic populations. With a basis of 119 "effective AOs" the errors introduced are essentially within the accuracy of the numerical integration. With a basis of 71 elements, the average and maximum errors in the atomic population are 0.007 and 0.016, respectively. In case of using only the minimal basis set of 37 elements, the errors increase but still one can see that the atomic populations can be reproduced within 0.1 electrons. Interestingly, these results appear to be almost independent of the original basis set.

It is worth to note that the use of the "effective AOs" as a numerical basis set is not restricted to the framework of QTAIM of disjunct atomic domains. In fact, in the framework of "fuzzy" atoms the "effective AOs" do not strictly form an orthogonal basis set but on the other hand they can be used to construct in a systematic manner numerical basis sets of increasing accuracy for atoms (or functional groups), that could be used in fully numerical[43] *ab initio* calculations. Of course, one should expect that in the case of correlated methods the size of the numerical basis set should be larger, for a better description of the virtual orbitals.

CONCLUSIONS

The "effective AOs" have been realized in the framework of Bader's QTAIM. This formalism can be used to retrieve from any type of calculation a proper set of orthogonal atomic basis functions. They form an orthonormalized set of numerical atomic orbitals, with occupation numbers that sum up to the respective QTAIM atomic populations. Importantly, only

a limited number of "effective AOs" exhibit significant occupation numbers, i.e., these atomic hybrids closely mimic the core and valence shells of the atom. In the case of hypervalent atoms, there appear additional hybrids with small but non-negligible occupation numbers.

We have shown that the MOs can be exactly expressed as a linear combination of this orthonormalized set of numerical atomic orbitals. Moreover, the Mulliken population analysis carried out on the basis of "effective AOs" exactly reproduces the original QTAIM atomic populations of the atoms. Approximate expansion of the MOs over a much reduced set of orthogonal atomic basis functions can also be accomplished to a very good accuracy with a SVD procedure.

Thus, this shows that there is nothing fundamentally inappropriate with a Hilbert-space based population analysis. The flaws of the classical Mulliken populations are rooted in the use of unsuitable atomic basis functions, not in its mathematical framework.[44]

APPENDIX A: EQUIVALENCE OF HILBERT-SPACE ANALYSIS IN THE BASIS OF EFFECTIVE AOs AND QTAIM ANALYSIS

By virtue of the definition (2), orbitals $\chi_\mu^A(\vec{r})$ differ from zero only in the atomic domain of atom A. They are orthonormalized

$$
\begin{aligned}
\langle \chi_\mu^A | \chi_\nu^A \rangle &= \left\langle \frac{1}{\sqrt{\lambda_\mu^A}} \sum_{i=1}^n U_{i\mu}^A \varphi_i^A \middle| \frac{1}{\sqrt{\lambda_\nu^A}} \sum_{j=1}^n U_{j\nu}^A \varphi_j^A \right\rangle \\
&= \frac{1}{\sqrt{\lambda_\mu^A \lambda_\nu^A}} \sum_{i,j=1}^n (\mathbf{U}^{A\dagger})_{\mu i} Q_{ij}^A U_{j\nu}^A \\
&= \frac{1}{\sqrt{\lambda_\mu^A \lambda_\nu^A}} \lambda_\mu^A \delta_{\mu\nu} = \delta_{\mu\nu} \qquad (A1)
\end{aligned}
$$

as a consequence of the eigenvalue equation (5). In fact, orbitals χ_μ^A represent the functions obtained by performing Löwdin's "canonic" orthogonalization of the functions φ_i^A. (Not to be confused with the usual Löwdin-orthogonalization performed by using matrix $\mathbf{S}^{-1/2}$.)

Relationship (6) can be trivially inverted, and one gets

$$\varphi_j^A = \sum_{\mu=1}^{n_A} U_{j\mu}^{A*} \sqrt{\lambda_\mu^A} \chi_\mu^A. \tag{A2}$$

Owing to this result and Eq. (4), the MOs can be written as linear combinations of the "effective AOs" of different atoms

$$\varphi_i = \sum_{A}^{N_{at}} \sum_{\mu=1}^{n_A} U_{i\mu}^{A*} \sqrt{\lambda_\mu^A} \chi_\mu^A. \tag{A3}$$

Thus, our analysis of the behaviour of the molecular orbitals in the different atomic domains resulted in a special LCAO expansion of the molecular orbitals. One can also introduce a continuous numbering of the basis orbitals, and write

$$\varphi_i = \sum_{\nu=1}^{m} C_{\nu i} \chi_\nu, \tag{A4}$$

where the overall dimension of the "effective AO" basis is

$$m = \sum_{A=1}^{N_{at}} n_A, \tag{A5}$$

and the orbitals χ_ν with $\nu \le n_1$ are attributed to atom with number $A = 1$, those with $n_1 + 1 \le \nu \le n_1 + n_2$ to atom $A = 2$, and so on, For the LCAO coefficients $C_{\nu i}$, one obviously has

$$C_{\nu i} = U_{i\mu}^{A*} \sqrt{\lambda_\mu^A}. \tag{A6}$$

If orbital χ_ν belongs to atom A, then subscripts μ and ν are trivially related as

$$\mu = \nu - \sum_{B=1}^{A-1} n_B. \tag{A7}$$

Owing to the orthogonality relationship (A1) and the disjunct character of the atomic basins, the orbitals χ_ν, $\nu = 1, 2, \ldots, m$ form an orthonormalized basis. Therefore, there are no overlap populations, and Mulliken's net and gross populations coincide. Thus, the Hilbert space or LCAO population of atom A in terms of the "effective atomic orbitals" is given by the sum of the diagonal density matrix elements for the orbitals belonging to that atom

$$Q_A^{LCAO} = \sum_{\nu \in A} D_{\nu\nu} = 2 \sum_{i=1}^{n} \sum_{\nu \in A} |C_{\nu i}|^2 = 2 \sum_{i=1}^{n} \sum_{\mu=1}^{n_A} \left| U_{i\mu}^{A*} \sqrt{\lambda_\mu} \right|^2$$

$$= 2 \sum_{i=1}^{n} \sum_{\mu=1}^{n_A} U_{i\mu}^{A} U_{i\mu}^{A*} \lambda_\mu = \sum_{i=1}^{n} \sum_{\mu,\nu=1}^{n_A} U_{i\mu}^{A} U_{i\nu}^{A*} \lambda_\mu \delta_{\mu\nu}$$

$$= \sum_{i=1}^{n} \sum_{\mu,\nu=1}^{n} U_{i\mu}^{A} \Lambda_{\mu\nu} (U^{A\dagger})_{\nu i}$$

$$= 2 \sum_{i=1}^{n} (U^A \Lambda^A U^{A\dagger})_{ii} = 2 \sum_{i=1}^{n} Q_{ii}^A, \tag{A8}$$

where the inverse of Eq. (5) has been utilized. (The summation limit for μ, ν was increased from n_A to n because that meant only adding terms containing factors $\lambda_\mu = 0$.)

The AIM population of atom A is given by

$$Q_A^{AIM} = \int_{\Omega_A} \rho(\vec{r}) dv = 2 \int_{\Omega_A} \sum_{i=1}^{n} |\varphi_i(\vec{r})|^2 dv$$

$$= 2 \sum_{i=1}^{n} \int_{\Omega_A} |\varphi_i(\vec{r})|^2 dv = 2 \sum_{i=1}^{n} Q_{ii}^A. \tag{A9}$$

Comparing Eqs. (A8) and (A9), we see that

$$Q_A^{LCAO} = Q_A^{AIM}, \tag{A10}$$

i.e., the atomic population calculated by the Hilbert-space analysis in the basis of the "effective AOs" is equal to that obtained by the 3D AIM analysis.

APPENDIX B: APPROXIMATE EXPANSION OF THE MOs IN THE BASIS OF "EFFECTIVE AOs"

Let us consider those "effective AOs" χ_μ^A, $A = 1, 2, \ldots N_{at}$ which meet some criterion $\lambda_\mu^A \ge t \ge 0$. Let their effective number be n_{eff}. In order to get an (approximate) expansion of the MOs, we should take enough "effective AOs," so it must be $n_{eff} \ge n$.

Now we build the rectangular $n_{eff} \times n$ matrix $\mathbf{Z}$, with elements

$$Z_{\mu i} = \langle \chi_\mu | \varphi_i \rangle. \tag{B1}$$

With the SVD, the rectangular matrix is transformed as

$$\mathbf{U}^\dagger \mathbf{Z} \mathbf{V} = \Xi, \tag{B2}$$

where $\mathbf{U}$ and $\mathbf{V}$ are unitary matrices of dimension $n_{eff} \times n_{eff}$ and $n \times n$, respectively, and Ξ is a rectangular diagonal matrix containing the singular values ξ_i of $\mathbf{Z}$. From the definition of matrix $\mathbf{Z}$ and using Eq. (B2), one can write

$$\sum_{\mu}^{n_{eff}} \sum_{j}^{n} U_{\mu i}^* \langle \chi_\mu | \varphi_j \rangle V_{ji} = \xi_i, \tag{B3}$$

that is, the singular value ξ_i is the overlap between the function

$$\psi_i = \sum_{\mu=1}^{n_{eff}} U_{\mu i} \chi_\mu, \tag{B4}$$

which is a linear combination of the "effective AOs" and

$$\varphi_i' = \sum_{\mu=1}^{n} V_{ji} \varphi_j, \tag{B5}$$

representing a molecular orbital after performing a unitary transformation with the matrix $\mathbf{V}$.

If $\xi_i = 1$, the two functions have an overlap equal one, ψ_i and φ_i' are essentially ("almost everywhere") equal to each other, and one can write

$$\varphi_i' = \sum_{\mu}^{n_{eff}} U_{\mu i} \chi_\mu. \tag{B6}$$

That is, the columns of the unitary matrix $\mathbf{U}$ contain the LCAO coefficients of each rotated MO in the orthogonal basis of "effective AOs."

If ξ_i is close to, but not exactly equal one, Eq. (B6) represents an approximation to the (rotated) molecular orbital. Under these circumstances, Eq. (A10) is no longer strictly fulfilled but, since these approximated MOs form an orthonormalized set, the number of electrons is conserved. (The "effective AOs" χ_μ are orthonormalized and matrix **U** is unitary, so we have n orthonormalized approximate MOs φ_i', so the closed shell determinant wavefunction constructed by their use carries exactly as many electrons as the original wavefunction containing the orbitals φ_i.)

Finally, the SVD method can also be applied in practice beyond doubly occupied orbitals. First of all, in order to reduce the dimensionality of the problem, it is worth to consider the natural orbital representation, ψ_i^{nat}, and thus perform the pairing between the set of "effective AOs" on one side and a reduced number of natural orbitals, n_{occ} (chosen again by an occupation number criterion). After the SVD process, each rotated natural orbital is identified with a linear combination of "effective AOs," as in the previous case

$$\sum_j^{n_{occ}} V_{ji} \psi_j^{nat} \cong \sum_\mu^{n_{eff}} U_{\mu i} \chi_\mu. \tag{B7}$$

However, now it is more convenient to have the connection directly with the original set of natural orbitals, where the density matrix is diagonal. Multiplying Eq. (B7) by the matrix element $(\mathbf{V}^\dagger)_{ik} = V_{ki}^*$, and summing up over i, one gets

$$\psi_k^{nat} \cong \sum_\mu^{n_{eff}} \sum_i^{n_{occ}} U_{\mu i} (\mathbf{V}^\dagger)_{ik} \chi_\mu = \sum_\mu^{n_{eff}} W_{\mu k} \chi_\mu, \tag{B8}$$

where now the columns of matrix $\mathbf{W} = \mathbf{U}\mathbf{V}^\dagger$ gather the (approximate) LCAO coefficients of the original subset of natural molecular orbitals over the "effective AO" basis.

[1] R. S. Mulliken, J. Chem. Phys. **23**, 1833 (1955).
[2] P.-O. Löwdin, Adv. Quantum Chem. **5**, 185 (1970).
[3] F. L. Hirshfeld, Theor. Chim. Acta **44**, 129 (1977).
[4] R. F. W. Bader, *Atoms in Molecules: A Quantum Theory* (Oxford University Press, Oxford, 1990).
[5] I. Mayer and P. Salvador, Chem. Phys. Lett. **383**, 368 (2004).
[6] T. C. Lillestolen and R. J. Wheatley, Chem. Commun. **45**, 5909 (2008).
[7] P. Bultinck, C. Van Alsenoy, P. W. Ayers, and R. Carbó-Dorca, J. Chem. Phys. **126**, 144111 (2007).
[8] T. C. Lillestolen and R. J. Wheatley, Chem. Commun. **2008**, 5909.
[9] P. Bultinck, D. L. Cooper, and D. V. Neck, Phys. Chem. Chem. Phys. **11**, 3424 (2009).
[10] T. C. Lillestolen and R. J. Wheatley, J. Chem. Phys. **131**, 144101 (2009).
[11] I. Mayer, J. Comput. Chem. **28**, 204 (2007).
[12] J. Baker, Theor. Chim. Acta **68**, 221 (1985).
[13] E. R. Davidson, J. Chem. Phys. **46**, 3320 (1967).
[14] R. Heinzmann and R. Ahlrichs, Theor. Chem. Acc. **42**, 33 (1976).
[15] J. G. Ángyán, M. Loos, and I. Mayer, J. Phys. Chem. **98**, 5244 (1994).
[16] E. Ramos-Cordoba, E. Matito, P. Salvador, and I. Mayer, Phys. Chem. Chem. Phys. **14**, 15291 (2012).
[17] P. Salvador and I. Mayer, J. Phys. Chem. **120**, 5046 (2004).
[18] M. A. Blanco, A. Martín Pendás, and E. Francisco, J. Chem. Theory Comput. **1**, 1096 (2005).
[19] E. Francisco, A. M. Pendás, and M. A. Blanco, J. Chem. Theory Comput. **2**, 90 (2006).
[20] E. Ramos-Cordoba, E. Matito, I. Mayer, and P. Salvador, J. Chem. Theory Comput. **8**, 1270 (2012).
[21] M. D. Pedro Salvador and I. Mayer, J. Chem. Phys. **115**, 1153 (2001).
[22] D. R. Alcoba, L. Lain, A. Torre, and R. C. Bochicchio, Chem. Phys. Lett. **407**, 379 (2005).
[23] L. Li and R. G. Parr, J. Chem. Phys. **84**, 1704 (1986).
[24] D. R. Alcoba, L. Lain, A. Torre, and R. C. Bochicchio, J. Chem. Phys. **123**, 144113 (2005).
[25] D. Vanfleteren, D. V. Neck, P. Bultinck, P. W. Ayers, and M. Waroquier, J. Chem. Phys. **133**, 231103 (2010).
[26] R. McWeeny, Rev. Mod. Phys. **32**, 335 (1960).
[27] I. Mayer, Chem. Phys. Lett. **242**, 499 (1995).
[28] I. Mayer, J. Phys. Chem. **100**, 6249 (1996).
[29] I. Mayer, Can. J. Chem. **74**, 939 (1996).
[30] I. Mayer and P. Salvador, J. Chem. Phys. **130**, 234106 (2009).
[31] I. Mayer, I. Bakó, and A. Stirling, J. Phys. Chem. A **115**, 12733 (2011).
[32] R. Ponec, J. Math. Chem. **21**, 323 (1997).
[33] D. Tiana, E. Francisco, M. A. Blanco, P. Macchi, A. Sironi, and A. Martin Pendás, Phys. Chem. Chem. Phys. **13**, 5068 (2011).
[34] J. Cioslowski, Int. J. Quantum Chem. **38**, 015 (1990).
[35] X. Fradera, M. A. Austen, and R. F. Bader, J. Phys. Chem. A **103**, 304 (1999).
[36] P. Salvador and E. Ramos-Cordoba, APOST-3D, University of Girona, Spain, 2011.
[37] J. I. Rodríguez *et al.*, J. Comput. Chem. **30**, 1082–1092 (2009).
[38] M. J. Frisch *et al.*, GAUSSIAN 03, Revision C.02, Gaussian, Inc., Pittsburgh, PA, 2003.
[39] P. Karadakov, D. Cooper, T. Thorsteinsson, and J. Gerratt, "Modern Valence-Bond Description of the Mechanisms of Six-Electron Pericyclic Reactions," in *Quantum Systems in Chemistry and Physics. Volume 1: Basic Problems and Models Systems*, edited by A. Hernández-Laguna, J. Maruani, R. McWeeny and S. Wilson (Kluwer, Dordrecht, 2000) pp. 327–344.
[40] W. Sander, D. Grote, S. Kossmann, and F. Neese, J. Am. Chem. Soc. **130**, 4396 (2008).
[41] R. Ponec, E. Ramos-Cordoba, and P. Salvador, J. Phys. Chem. A **117**, 1975 (2013).
[42] E. Ramos-Cordoba and P. Salvador, "Local spin analysis of polyradicals" (unpublished).
[43] M. Zakharov, "Performance of numerical atom-centered basis sets in the ground-state correlated calculations of noncovalent interactions: Water and methane dimer cases," Int. J. Quantum Chem. (published online).
[44] I. Mayer, Chem. Phys. Lett. **97**, 270 (1983).